

# Release A CDR RID Report

Date Last Modified 10/27/95

Originator Chris Lynnes

Phone No 301-286-2260

Organization GSFC DAAC

E Mail Address lynnes@daac.gsfc.nasa.gov

Document Ingest Design

RID ID CDR 14

Review SDPS/CSMS

Originator Ref GD-CDR-CL-12

Priority 2

Section

Page EN-11

Figure Table

Category Name Ingest Design

Actionee ECS

Sub Category

Subject Document Pre-processing by Ingest

## Description of Problem or Suggestion:

Multi-file HTML documents will include URLs pointing to each other on the host where they were developed. The files must be modified so the URLs point to where the documents will be when they are in the document server. (Presumably, this is analogous to data reformatting).

## Originator's Recommendation

Since this reformatting is predictable and well-understood, it should be included as an ingest pre-processing method. (If not an ingest function, this should be allocated to the Document Server)

GSFC Response by:

GSFC Response Date

HAIS Response by: C. Gire

HAIS Schedule 9/6/95

HAIS R. E. LiLing Chao

HAIS Response Date 10/25/95

For HTML multi-document ingest, the Ingest Subsystem will perform validations such as syntax check, URL check, and metadata check prior to inserting the documents into the Document Data Server as required. Ingest will not modify the URLs.

In order for ECS to guarantee successful retrieval of HTML documents to be ingested into the ECS system, the following rules must be followed:

- 1) Only relative HTML links (i.e., no full path is specified) can be used to refer to other HTML documents in a multi-file document data set. This is the user's responsibility.
- 2) HTML documents of the same logical group have to be ingested with a single ingest request. This is the user's responsibility.
- 3) The file names of the HTML documents will be maintained by the Document Data Server. This is an ECS responsibility.

Status **Closed**

Date Closed **10/27/95**

Sponsor **Kobler**

\*\*\*\*\* Attachment if any \*\*\*\*\*